

REMARKS

Claims 1-4, 7-19, and 21-38 are pending in the application. Claims 1, 19, and 23-25 are currently amended.

The Examiner is thanked for granting the interview dated October 24, 2006.

In the Interview the issue of the term "scanned" was discussed in respect of claim 1, and the difference between the intermediate and final formats was also discussed, both in respect of claims 1 and 25. In light of the interview, Applicant adds the term "scanned" to claim 1, and provides further clarifications in the remarks.

Claim Rejections – 35 USC 103

In this section of the Office Action, claims 1-4, 7-19, and 21-38 were rejected under 35 USC 103(a) as being unpatentable over Ferrel et al., US006199082B1 in view of Fitzsimons et al, US006708189B1, and in further view of Blumberg et al, US 20030140315A1.

Favorable reconsideration of this rejection is respectfully requested since, as will be shown below, the above-amended claims are both novel and inventive over the prior art cited by the Examiner.

The present invention relates to a system which automatically carries out intelligent structure analysis of a scan or like existing passive digital version of an existing document. The scan or like version preserves the layout of the document, but is digitally inactive. That is to say it is electronically passive. There are no links that can be used or features that can be activated. The document is analyzed to decompose the structure of the document, in order to understand the separate parts of a document and the internal structures of those parts.

The system is then used to parse documents, using a marking system to mark the structure on the documents and then using the marking to republish the original document in a second format.

This is discussed in the present application, specifically at the first paragraph of page 8, wherein it is noted that "the present invention does not merely convert the data from one format to a different format, but instead is able to also provide at least a basic structure for organizing the data into the Web page."

The system of the present invention is detailed further in the specification, at page 11, third paragraph – page 17, wherein it is stated that, according to the present invention,

data which is in a digital format can be converted to a basic internal format, which is preferably XML, although substantially any other type of mark-up language could also be used. Next, intelligent structure analysis is performed, in order to recognize and define the structures and objects contained in the newspaper data, particularly with regard to each page of the newspaper.

More preferably, newspaper objects are classified as one of a plurality of specific entities, which optionally and most preferably are selected from the group consisting of an article, an advertisement, a picture which is not otherwise associated with an article or advertisement, and general data, which covers information that does not fit into one of the preceding categories.

Recognition of the hierarchy for newspaper objects is an important part of the process. The process of intelligent structure analysis enables the newspaper data to be converted to a series of objects, for more efficient search and retrieval through the Internet or other network. After the process of intelligent structure analysis has been completed, encoding of the object data is performed.

This results in a set of enhanced, structured files which combine the original image of the data with the text and encoding information. Each such file thus preferably maintains the visual aspects of the newspaper layout, while enabling far greater functionality to be available through the Web page version of the newspaper. The internal format data is published in a plurality of different final formats by a publication server.

According to the present invention, typically, a document in an original, existing format is divided into pages, articles, pictures, pictures with articles, articles with headlines, articles with headlines and a by-line, articles with headlines, by-line, a picture, and the like. The document could be a newspaper, a company financial report, or any kind of structured document. The final published format is an interactive document, typically for use as a Web page.

The automatic parsing of the present embodiments allows the structural subunits, the blocks, of the original document to be the interactive objects of the Web page. Thus a particular block, comprising a headline, by-line, text and a photograph, is identified automatically as belonging together. The block is defined as a single object, and the text, etc., as internal structure thereof. The parser classifies the headline using an XML label, the by-line likewise, the text likewise, and the picture likewise, and then the publisher knows that the entire object should be presented visually as it appears in

the original newspaper page. However it also knows how to set up the headline as a hypertext link to lead to a larger size version of the full text so that the text can be read. Likewise it knows how to set up the picture as a link that can be clicked on to lead to a larger view of the picture. The by-line may be treated separately to provide a link to other articles by the same author. The result is a fully interactive version of the original document.

In summary, the interactive version of a multi-element document is automatically arrived at by reverse engineering (parsing) of *a received* preexisting document in the original, existing format, leading to XML labeling of the document parts to represent blocks and internal structure of the blocks, thus generating an internal XML based format, representative of the original document.

The internal format is used as directions by a publisher to publish an interactive version of the document, in a final publication format.

Ferrel discloses a multimedia publishing system for creating online publications, wherein the contents and design aspects of the online publishing process are separated into distinct components before being transmitted to the customers' computers. Separation of the contents and the design aspects of the online publications facilitates efficient distribution of published documents by enabling the transmission of, for example, high-quality titles over low-speed communications links subject to loss of connectivity.

Fitzsimons discloses a computer file transfer system wherein a data file is received by an automated publishing system from a source, and the data file is converted for use within a presentation space of the automated publishing system. As discussed at column 12, line 60 – column 13, line 55 of the patent to Fitzsimons, the system converts a file, such as an MS Word document file or a digital image TIFF file, created under a first format, into a corresponding data file under a second format. This enables communication between computers which must exchange files having different data structures.

Blumberg discloses a system for on-line on-demand printing services that enable a user to interactively create and order print jobs via a computer network. With

Blumberg, a user is allowed to select finishing options for the print jobs, such as paper grain, front cover, etc, and view an on-line proof of the finished printed jobs.

For example, Blumberg describes in paragraph [0028]: " In one embodiment, intended for use over the Internet and referred to as an Internet print on demand service, the present invention enables a user to submit over the Internet his own document, and to select over the Internet various finishing options for printing. As used in the present specification, "finishing options" refer to physical attributes of a printed document *that are not related to the content of the document*. Examples of finishing options include type of paper (grain, weight, size and color), type of binding, types of front and back covers, type of tab inserts, type of lamination, type of folding and type of hole punching".

That it to say, Blumberg aims at providing an improved method for on-line finishing printing option selection and visualization of final printed documents, so as to provide an on-line user with the ability to influence and predict the "looks and feel" of the finished printed document (a book, a newspaper, etc), according to the finishing options selected by the user.

None of the cited documents either shows or suggests automatically analyzing a *received original, existing document*, such as a newspaper or a microfiche carried document, in such a reverse engineering process, in order to understand the separate parts of a document and the internal structures of those parts, as taught by the present invention.

Claim 1, as currently amended, defines a method for automatically publishing data in a final publication format, wherein the data is in the form of a newspaper having an original, existing format including a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having an internal structure, the method comprising: obtaining a *scanned* representation of the newspaper, the representation preserving the layout, from the representation automatically analyzing the newspaper to decompose the predetermined layout of each page of the newspaper in the original, existing format into the plurality of blocks, each block representing an object; converting each object to an internal publication format, the internal publication format identifying and preserving the internal structure of the blocks within the objects, the internal

publication format furthermore preserving the layout as a relationship between the objects; and rendering the internal publication format to incorporate the objects, the layout and respective internal structures in the final publication format.

As explained above, and defined by claim 1, the present invention introduces the novel and inventive idea of a method where an *obtained scanned representation* of an *existing* publication, such as a newspaper is automatically analyzed, in order to understand the separate parts of the published document and the internal structures of those parts, and generate an internal format representative of the structure and objects of the originally scanned document, as described in 24, page 11: "Next, XML distiller module 18 preferably performs *intelligent structure analysis*, in order to be able to recognize and define the structures and objects contained in the newspaper data, particularly with regard to each page of the newspaper. Examples of such structures and objects include, but are not limited to, articles, advertisements, titles, and so forth".

In contrast, Ferrel introduces a publication system wherein *the publication process*, carried out for the initial generation of the published document, is divided into a content aspect and a design aspect. However, Ferral does not describe or even hint at the idea of a reverse engineering process where a *received representation of an existing* publication is analyzed, to decompose the predetermined layout of each page of the newspaper in the original, existing format into the plurality of blocks etc, as taught by the present invention and defined by claim 1.

As described hereinabove, Fitzsimons discloses a computer file transfer system. Fitzsimons also falls short of disclosing or even hinting at the idea of a reverse engineering process where a *received scanned representation* of an existing *newspaper* is analyzed, to decompose the predetermined layout of each page of the newspaper in the original, existing format into the plurality of blocks etc, as taught by the present invention and defined by claim 1.

As described hereinabove, Blumberg aims at providing an improved method for on-line finishing printing option selection and visualization of final printed documents, so as to provide an on-line user with the ability to influence and predict the "looks and feel" of the finished printed document (a book, a newspaper, etc), according to the finishing options selected by the user.

For example, Blumberg describes in paragraph [0028]: " In one embodiment, intended for use over the Internet and referred to as an Internet print on demand service, the present invention enables a user to submit over the Internet his own document, and to select over the Internet various finishing options for printing. As used in the present specification, "finishing options" refer to physical attributes of a printed document *that are not related to the content of the document*. Examples of finishing options include type of paper (grain, weight, size and color), type of binding, types of front and back covers, type of tab inserts, type of lamination, type of folding and type of hole punching".

However, Blumberg also falls short of describing or even hinting at the idea of a reverse engineering process where a *received scanned representation of an existing* publication is analyzed, to decompose the predetermined layout of each page of the newspaper in the original, existing format into the plurality of blocks etc, as taught by the present invention and defined by claim 1. Claim 1 further defines that the substructure of the blocks preserves item substructure and the relationship between the blocks preserves page format structure. Neither of these features is hinted at in Blumberg, and certainly not the combination.

It is thus believed that claim 1 is both novel and inventive over the prior art and respectfully maintained that the claim should be allowed.

Claim 19, as currently amended, defines a system for automatically publishing received data of a pre-existing newspaper, the newspaper in a computerized format, the system comprising:

- (a) at least one source of newspaper data, the source preserving an original, structure of the newspaper, the computerized format comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having an internal structure;
- (b) a mark-up language distiller module for converting the newspaper from the original format to a mark-up language format, wherein the mark-up language distiller module automatically analyzes the newspaper data in the original, existing digital format to decompose the newspaper data into the plurality of blocks, each block with the internal structure representing an independent data object, each object having content and at least one attribute of the data, such that each object is converted to the mark-up language format, the markup language format further analyzing and preserving the structure; and
- (c) a publisher server for converting the data from the mark-up language format to a final publication

format, the final publication format incorporating, as objects, the blocks with the internal structure, and further defining interrelationships with the objects.

As explained above, and defined by claim 19, the present invention introduces the novel and inventive idea of a system for automatically publishing *received* data of *a pre-existing* newspaper, the newspaper having a computerized format. The system includes at least one source of *a newspaper* preserving the original, *existing* computerized format of the newspaper, the format comprising a plurality of pages etc.

As explained above, neither Ferrel nor Fitzsimons nor Blumberg discloses or even hints at the idea of a system automatically publishing *received* data of *a pre-existing* newspaper in a computerized format, the system having at least one source of *a newspaper* preserving the original, *existing* digital format of the newspaper, as taught by the present invention. Claim 19 further defines that the substructure of the blocks preserves item substructure and the relationship between the blocks preserves page format structure. Neither of these features is hinted at in Blumberg, and certainly not the combination.

It is thus believed that claim 19 is both novel and inventive over the prior art and respectfully maintained that the claim should be allowed.

Claim 23, as currently amended, defines a method for automatically publishing data of a preexisting document in a final publication format, wherein the data is received in an original, existing format comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having an internal structure the method comprising: obtaining a representation of the preexisting document, the representation preserving the layout; from the representation, automatically analyzing the preexisting document to decompose the data received in the original, existing format into a plurality of objects, each object corresponding to one of the blocks; preparing a list of text and/or graphic elements for each object; determining properties of each element, including determining visibility and overlap characteristics for each graphic element within the object; recognizing structural layout properties of the data in an original format; converting each object to an internal publication format; and rendering the internal publication format in the final publication format, the final publication format

presenting the blocks as the independently standing objects incorporating the internal structure and the layout.

As explained above, neither Ferrel nor Fitzsimons nor Blumberg discloses or even hints at the idea of a method of automatically publishing *received* data of a *pre-existing* document in a computerized format, the method including automatically analyzing an obtained representation of the pre-existing document to decompose the data *received* in the original, existing format into a plurality of objects etc, as taught by the present invention. Claim 23 further defines that the substructure of the blocks preserves item substructure and the relationship between the blocks preserves page format structure. Neither of these features is hinted at in Blumberg, and certainly not the combination.

It is thus believed that claim 23 is both novel and inventive over the prior art and respectfully maintained that the claim should be allowed.

Claim 24, as currently amended, defines a method for automatically publishing received data in a final publication format, the data having an original, existing format comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, the method comprising: obtaining a representation of the received data having an original existing format, the representation preserving the layout; from the representation, automatically analyzing the pages to decompose the pages of the data received in the original, existing format into a plurality of objects, the objects corresponding to the blocks; preparing a list of text and/or graphic elements for each object; determining properties of each element, including determining a special characteristic for each text element; recognizing structural layout properties of the data in an original format; converting each object to an internal publication format; and rendering the internal publication format in the final publication format such as to include the recognized structure in the objects and to include the layout.

As explained above, neither Ferrel nor Fitzsimons nor Blumberg discloses or even hints at the idea of a method which comprises automatically analyzing a representation to decompose the data *received* in the original, existing format into a plurality of objects, etc, as taught by the present invention. Claim 24 further defines that the substructure of the blocks preserves item substructure and the relationship

between the blocks preserves page format structure. Neither of these features is hinted at in Blumberg, and certainly not the combination.

It is thus believed that claim 24 is both novel and inventive over the prior art and respectfully maintained that the claim should be allowed.

Claim 25, as currently amended, defines a method for automatically publishing received data in a final publication format, wherein the data is in the form of a newspaper, the newspaper having an original, existing format comprising a plurality of pages, each page having a predetermined layout comprising a plurality of independently standing data blocks, each block having structural layout properties, the method comprising: obtaining a representation of said newspaper, said representation preserving said layout; from said representation, automatically analyzing the newspaper to decompose the data received in the original, existing format into a plurality of objects, said objects corresponding to said independently standing blocks; preparing a list of text and/or graphic elements for each object; determining properties of each element; recognizing said structural layout properties of the data in an original format; determining each text segment for each object; building a text block from a plurality of aligned text segments; converting each object to an internal publication format; and rendering said internal publication format in the final publication format to comprise said blocks as objects incorporating said structural layout properties, and to include said layout.

As explained in further detail hereinabove, and defined by claim 25, the present invention introduces the novel and inventive idea of a method for automatically publishing automatically publishing data *of a newspaper* in a final publication format. The method includes automatically analyzing *a representation of the newspaper*, to decompose the data *received* in the original, existing format into a plurality of objects. As explained above, neither Ferrel nor Fitzsimons nor Blumberg discloses or even hints at the idea of a method which comprises automatically analyzing the data to decompose the data *received* in the original, existing format into a plurality of objects, etc, as taught by the present invention. Claim 25 further defines that the substructure of the blocks preserves item substructure and the relationship between the blocks preserves page format structure. Neither of these features is hinted at in Blumberg, and certainly not the combination.

It is thus believed that claim 25 is both novel and inventive over the prior art and respectfully maintained that the claim should be allowed.

The remaining claims mentioned in this section of the Office Action are believed to be allowable as being dependent on an allowable main claim.

All of the matters raised by the Examiner have been dealt with and are believed to have been overcome.

In view of the foregoing, it is respectfully submitted that all the claims now pending in the application are allowable.

An early Notice of Allowance is therefore respectfully requested.

Respectfully submitted,



Martin D. Moynihan
Registration No. 40,338

Date: November 20, 2006

Encl.:

Petition for Extension of Time (2 Months)